AUTOMATED DIAGNOSIS OF BRAIN TUMOUR IN MRI IMAGES USING DEEP LEARNING TECHNIQUES

Sri hari gupta K
dept. of computer science and engineering, Lovely Professional University, Phagwara, Punjab.
sriharigupta55@gmail.com

Jaspreet Kaur (asst. professor)
dept. of computer science and engineering, Lovely Professional University, Phagwara, Punjab.
jaspreet.kaur144@gmail.com

Abstract

Brain is one of the most vital organ in human nervous system. Brain contains most lacks of nerves connected and communicating each other. An unusual growth or increase of nerves in the brain, which may affect the actual working of the brain, is called a Brain tumour. In the modern day the brain tumour was detected by traditional way of laborious and traditional approach. However, this lab analysis was too time consuming, which may lead to the increase the severity of the tumour cells. To address this issue, automated brain tumour detection techniques was introduced, which will help the early detection of tumour cells and helps for the timely medicating. In this study, The Cancer Imaging Archive was utilized as a dataset. The pre-processing and data augmentation techniques have been applied for training purposes using various deep learning (DL) models. The optimizers were used to determine stronger dataset to train the model. The pre-processed dataset were passed to the model to train on it. The deep learning (DL) model Convolutional Neural Network is used with the various pre-trained models like ResNet50, Xception, VGG19, MobileNetV2, InceptionV3. The accuracy of the different models was validated and tested to obtain the model which gives the good accuracy. This research primarily focusing to help the medical professionals in their efforts to detect brain tumour through the use of imaging techniques.

Keywords: *Pre-processing, CNN, ResNet50, Xception, VGG19, MobileNetV2, InceptionsV3.*

1. Introduction

Human brain controls every aspect of human actions ranging from physical activities, thoughts to emotions. Brain is functioning by sending and receiving the signals to all over the body via neurons. Brain and Spinal cord connected each other and forming central nervous system (CNS). Brain tumour is the growth of the abnormal cells in the brain. The brains anatomy is very complex, each and every part of the brain is responsible of different functions. Brain tumour can develop in any part of the brain, including the protective lining and the skull base, the brainstem, the sinuses and the nasal cavities. In this recent years, brain tumours are the deadliest disease with highest death rate. The right strategy for knowing brain tumours and its stages is crucial to both prevention and carrying out the necessary therapeutic measures. In view of this, medics frequently employ magnetic resonance imaging, or MRI, to examine brain tumours [1].

Detecting the brain tumour is big task, scanning the brain to locate the tumour cells. There are many different scans such as Magnetic Resonance Images (MRI), Computed Tomography(CT), Biopsy, Angiography. In this project Magnetic Resonance Image (MRI) dataset were used. MRI images are sequenced as T1 weighted (T1W), T2 weighted (T2W), Fluid attenuated, Fluid Attenuated Inversion Recovery (FLAIR), Proton Density. In this project the FLAIR MRI image dataset is used to train the model. Advantages of using FLAIR MRI images is, it selectively suppresses the fluid signals such as cerebrospinal fluid (CSF), it will differentiate different leisure and tissue. However, the detection in traditional method takes very long time and it will increase the severity in that time.

With the help of the modern computer and latest technology brain tumour can be detected at early stage and correct medications will be provided at the right time. The FLAIR MRI images were pre-processed to ensure that the creating model should receive correct format to train without any interruption. The shape of the dataset should be in same size for avoiding the errors. Normalizing the data, removing the noise in the data is done in pre-processing state.

Data augmentation is a technique which is used to create modified datasets by using the given dataset. With this technique we can ensure that we can create a dataset without losing the features. Data augmentation is a technique which will rotate, shear, scale, flip the dataset and saving it as the new inputs. This technique will be effective when there is shorter in dataset. The augmented data were visualized and see how well the dataset was spread over the target, information of the datasets.

The Convolutional Neural Network (CNN) takes augmented images as input and train the model. The pretrained model were included and the dataset passed to it. Resnet50, Imagenetv3, Xception, VGG19, MobileNetv2 pretrained models were included. The last layer of this models were cut down and the output of is passed to the network for the classification. The output is tested with the metrics such as Precision, Recall, Specificity, F1 Score, Accuracy, Confusion matrix. The calculated metrics shows that the models performance over the dataset.

2. Literature Review

Brain tumour detection in MRI image is quite a bigger challenge in medical field. Most of the related work in the field of Brain tumour involves the data pre-processing, feature extraction, detection of tumour. The existing models were built in Deep learning (DL) and Convolutional Neural Network (CNN).

Ramtekkar et al. (2023) [1] proposes a method for tumour detection model. The authors used Kaggle dataset. The model involves pre-processing, filteration, segmentation, feature extraction, optimization, convolutional neural network (CNN). In pre-processing, noise in the MRI images were removed by converting the images into grayscale, and filter it out by filtering methods. The MRI images are segmented into non-overlapping-sections. To identify the boundaries thresholding and histogram based segmentation process is used. Feature extraction is done by contrast, correlation, homogeneity, entropy, energy is used. Optimization techniques like Particle swarm optimization (PSO), Genetic Algorithm (GA), Gray wolf optimization (GWO) is used to select the fittest image feature. Then, the images are passed to the CNN model where the model identifies the tumour in the image and classify accordingly. Finally, the model obtained the accuracy of 98.9%.

Kumar et al. (2023) [2] suggests a different approach over brain tumour classification. The authors used cancer imaging archives datasets. The model is comprised of pre-processing, segmentation, feature extraction and classification. In pre-processing the data are reshaped to organize it in same shape as 64 X 64. Then the grayscale and thresholding were utilised to boost the tumour region's intensity. In segmentation, instead of using tradition approach the authors used dual-tree wavelet transformation (DTWT) and used low pass filter. For classification, using pre-trained model and transfer learning. The Soft-max classifier is used for classification and probabilistic role in transfer training. UNet, AlexNet, VGG16, ResNet50 models were used. The ResNet 50 model is able to classify the malignant and benign class with 98.4% and 99.3% .

Nassar et al. (2024) [3] proposed a strategy that is for brain tumour classification technique in hybrid DL techniques. T1W-CE MRI dataset were used which includes meningioma, glioma, and pituitary MR image. Data augmentation technique were used only to training techniques and for testing the original dataset is used without any pre-processing for the originality. The image was resized so that it fit to the CNN. Here five altered fine-tuned pretrained models were used such as GoogleNet, AlexNet, ShuffleNet, SqueezeNet, and NASNET-Mobile to distinguish their performance in classifying different categories of brain trumors. The model evaluated using overall accuracy, f1-score and confusion matrix. The NASNet model yielded the accuracy of 97.50%.

Razzaq et al. (2024) [4] the authors proposed optimized linear support vector network (oLSVN) based brain tumour detection using MRI images. The data set used is comprised of 5710 images divided into testing and training datasets. The proposed methodology consists of pre-processing, feature extraction, hand crafted features, high quality feature selection, hybrid feature (HF) selection, Bag of Feature(Bof). The metrics like Precision, Recall, Specificity, Sensitivity, F1-score, Mathews Correlation Coefficient (MCC) is calculated to measure the model performance. The proposed model produces the accuracy of 87.4% with an execution time of 243.6 sec.

Haq et al. (2023) [5] suggest MRI based approach for efficient categorising of brain tumour. The researchers utilized the Multimodal Brain Tumour Image Segmentation Benchmark (BraTS) brain tumour 2018 dataset as an input. The proposed model consists of pre-processing, data augmentation, segmentation process, Deep convolutional neural network (DCNN). Denoising techniques were used to eliminate Gaussian noise, Salt and Pepper noise, and Speckle noise from the MRI images. Filter like Wiener and median are used to minimize the mean square errors. The integrated method focussed on fuzzy based and brainstorm based optimization technique was used for extracting. Leaky ReLU (LReLU), Soft-max were used as a activation functions. The model runs successfully with the accuracy of 96.5% in BraTS 2018 dataset.

P.Ramya et al. (2021), [6] researchers utilised deep super learning classification and cluster ensemble for splitting malignant cells. The model they created consists of Image registration, pre-processing, clustering methods like K-means clustering, Gaussian Mixture Model, Fuzzy based Clustering, SOM and Cluster ensemble. Firstly, the unprocessed set of data T1, T2, and FLAIR MRI of test images is pre-processed using Laplacian Cellular Automata Filtering. K-means, fuzzy, and SOM segment the registered image. The clustered matrix is grouped for ensemble to form final segmented image. This method added advantages to improve the classification. The proposed work producing good results with the accuracy of 96.3% and 95.8%.

Chanu and Thongam (2021), [7] the authors stated the DL methods to detect the brain tumour from the MRI images. The proposed techniques are 2D-CNN, Image acquisition, Image pre-processing, cropping, Image rotation, Noise filtering, Image scaling, and Image augmentation. Here the unpre-processed images were processed to single format of image as PNG. Noise filtering is defined to be a random dissimilarities of image intensity which is the grains in the image. The Convolutional Neural network was built with the convolutional layers, non-linear units like Rectified linear unit (ReLU), Max min average pooling were used, flattening layer, fully connected layers are used. The model acquires the accuracy of 97%.

Kalyani et al. (2023) [8] proposes the DL techniques for analysis of brain tumour on MRI brain tumour images. The techniques used in this model is Faster R-CNN, YOLOV3, YOLO-tiny, object detection. The dataset which is used for this model is Microsoft common objects in context (MSCOCO) datasets. YOLOV3 in this model uses Darknet-53 architecture which have many convolutional layers. Normalization of the dataset is applied through batch normalization. YOLOV3 – tiny is an another variant of YOLOV3 with a reduced depth of CNN. Moreover, YOLOv3 detector units has been behave like recognition unit. The proposed model gives the accuracy of 93.14%.

Rammurthy D et al. (2020) [9] describes a brain tumour detection by using the optimized technique Whale Harris Hawks Optimization (WHHO). In the proposed approach, the theories of rough sets and automata with cells are used to segment data. Cancer size, the local optical oriented pattern (LOOP), the mean, the variance, and kurtosis belong to the properties of the image which are extracted from the segments. The model is trained using the Deep Convolutional Neural Network (DCNN) and the Whale Optimising Technique (WHHO). The correctness of the algorithm is 81.6%.

Asif et al. (2023) [10] has produced an enhanced DL method for multi-class classification using transfer learning. Xception, DenseNet201, DenseNet121, ResNet152V2, and InceptionResNetV2 are the models chosen for the paper. The Figshare benchmark dataset is used to assess the performance. The MRI data were pre-processed and cropped to ensure the maximum accuracy. Also, the hyperparameter and optimization procedures are implemented for attaining better and robust model. The model with the Xception produces the higher accuracy of 99.67%.

Amin et al. (2018) [15] has stated that different approach for detecting the tumour by the feature fusion and machine learning. Here the author used the benchmarked dataset of BRATS datasets in different years (2012 – 2015). For calculating the metrics of the model precision, recall, accuracy, specificity is used. To segment the tumour cells in the MRI image is done by pixel based experimental results. Gabor wavelet transform and the histogram is employed to extract the features. Classification of Brain MRI images into benign and malignant is done by using the model Random Forest classifier. By employing this model with the added segmentation and feature extraction it reached the accuracy of 98%.

Anil et al. (2019) [12] used a transfer learning approach to detect the tumour cells in brain. The author used VGG16, VGG19, Alexnet deep learning model to train on the brain MRI datasets. In pre-trained models only the top layers were included and the output of the top layers is transferred to the fine tune layers. Among these pre-trained model, the VGG19 model gave the highest accuracy of 95.78%.

Sajid et al. (2019) [16] has produced the innovative approach for detecting the brain tumour by using the DL models with MRI datasets. The author used BRATS 2013 brain tumour datasets. The proposed model deals with the major problem overfitting by means of reducing it. For that the datasets were pre-processed, batch normalization, regularization is done before passing to the model for training. The dataset is trained under the deep learning models feed forward network, CNN and post processing steps. By utilizing this architecture, the model reached the milestone of 91% accuracy.

Kokila et al. (2021) [18] has shown an approach to identification and classification of tumour. To train the deep learning models Kaggle dataset repository of cancer Imaging is used. The CNN based classification is done by multi-task classification. The identification of brain tumour location is done by using CNN model segmentation. This proposed model achieved accuracy of 89%.

3. Proposed Architecture: -

Brain tumour detection by using the artificial intelligence in the early stage will help to medicate the patients at the correct time. For this the proposed work with the architecture of the deep learning techniques and the image classification techniques will help the pharmacist to detect early. In this architecture the deep learning model includes CNN, ResNet50, Xception, Inception, VGG19, MobileNetV2 were employed and trained. The detection of brain tumour is done is following steps.

The contribution of the proposed research as mentioned below:

The dataset utilized in this model has been explained in Section 3.1. Preprocessing the input data in 3.2 minimizes data redundancy. The data's tumour segmentation is separated using mask values in section 3.3. Following that, the MRI images' characteristics were extracted and presented in section 3.4. The extracted features have been enhanced and visualized in sections 3.5 and 3.6. The deep learning methods utilized in this investigation (CNN, ResNet50, Xception, InceptionV2, VGG19, and MobileNetV2) were then described in section 3.7. Section 3.8 provides an explanation of the findings.

3.1. Dataset: -

The Cancer Imaging Archive (TCIA) generated the set of images that was employed in the present study. The MR images and a locally produced FLAIR (Fluid Attenuated Inversion Recovery) abnormality segmentation mask have been incorporated in this dataset. It includes 110 people with cancer that possess at least fluid attenuated inversion recovery strand and genomic cluster data available from The Cancer Genome Atlas (TCGA) lower-grade glioma tumour collection. It includes images in the Tag Image File Format (TIFF) format, each one of which includes three channels. There are three distinct sequence in each of the 3 order channels of the 101 image dataset: pre-contrast, FLAIR, and post-contrast. 9 instances are lacking its post-contrast sequence, whereas 6 instances lacked with the pre-contrast sequence. The FLAIR images are then added to the three channels so as to recoup them for this lost segment. This dataset is organised in 110 folders and is structured and named in accordance with the case ID. Here, the model gets instruction using the FLAIR images. The link for the dataset is https://www.kaggle.com/datasets/mateuszbuda/lgg-mri-segmentation.

3.2. Pre-processing

This technique refers to the data integrity, cleaning, transforming, reduction, normalizing, discretization the data. The resizing technique is used to resize or converted the image into the same size. To ensure the images are in the same size so that the model will train correctly. Image transformation leads to rotating, scaling, shearing, flipping the images to create a dataset with different information without losing the originality. Discretization involves separating the continuous values to the discrete values. Normalization is a technique where the intensity levels where scaled to common range of 0 to 1, so that it is easy to visualize and train the model. Pre-processing is the cleaning the data that means that to remove the noise in the image. There are many pre known noises like Gaussian noise, Rayleigh noise, Exponential noise, Uniform noise, Salt and pepper noise, periodic noise etc. With these help of pre known noises the noise in the image was removed and it produce the approximately noise free image. Feature selection is the important process that the selecting a relevant feature or subset of the requirement from the image, so that the training can be done easily. Sampling the subset from the dataset helps to reduce the size of the dataset with preserving the important information. Here in the pre-processing of the dataset was done by the normalization of the data, sampling the data. It helped the model to overcome the data redundancy.

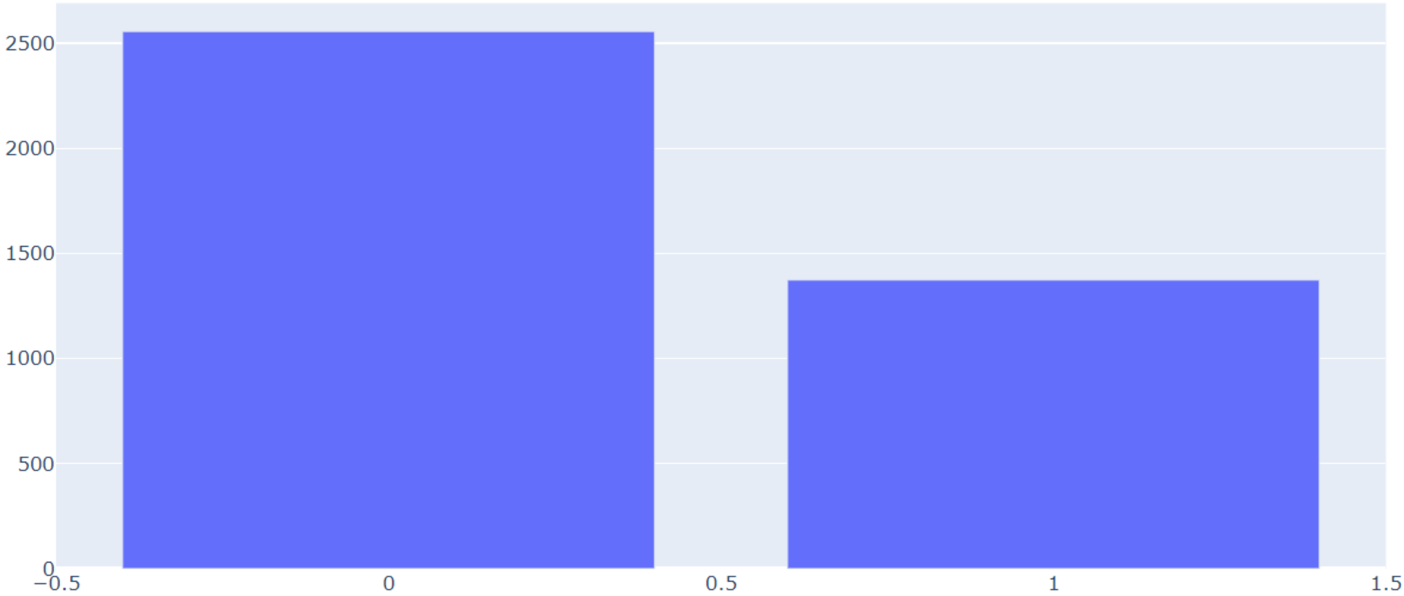


Figure 1. The above chart shows the distribution of the datasets.

3.3. Segmentation

The MRI image segmentation is very difficult because the many MRI images will be produced at the time of scanning and it made the task tougher to the physician to divide the images. At the time of taking MRI images, the images will be overlapped and it was confusing to view. So the task is to separate or partition the MRI images without any overlapping and in good condition. Segmentation is a task to partition the digital pictures into discrete group of pixels. In The Cancer Imaging Archive (TCIA) dataset the images T1 weighted, T2 weighted, Fluid Attenuated Inversion Recovery (FLAIR) images were grouped for 110 patients and stored is folder format. Here we are using Fluid Attenuated Inversion Recovery (FLAIR) images to train the model. The FLAIR images were separated from the group and the mask image of the image was separated. Mask image is a image which only contains the tumour cells. This separation will help the model to train correctly and majorly helps to test the model by visualizing the tumour cells that were predicted and in the mask cell. The FLAIR image and the Mask images were given below.

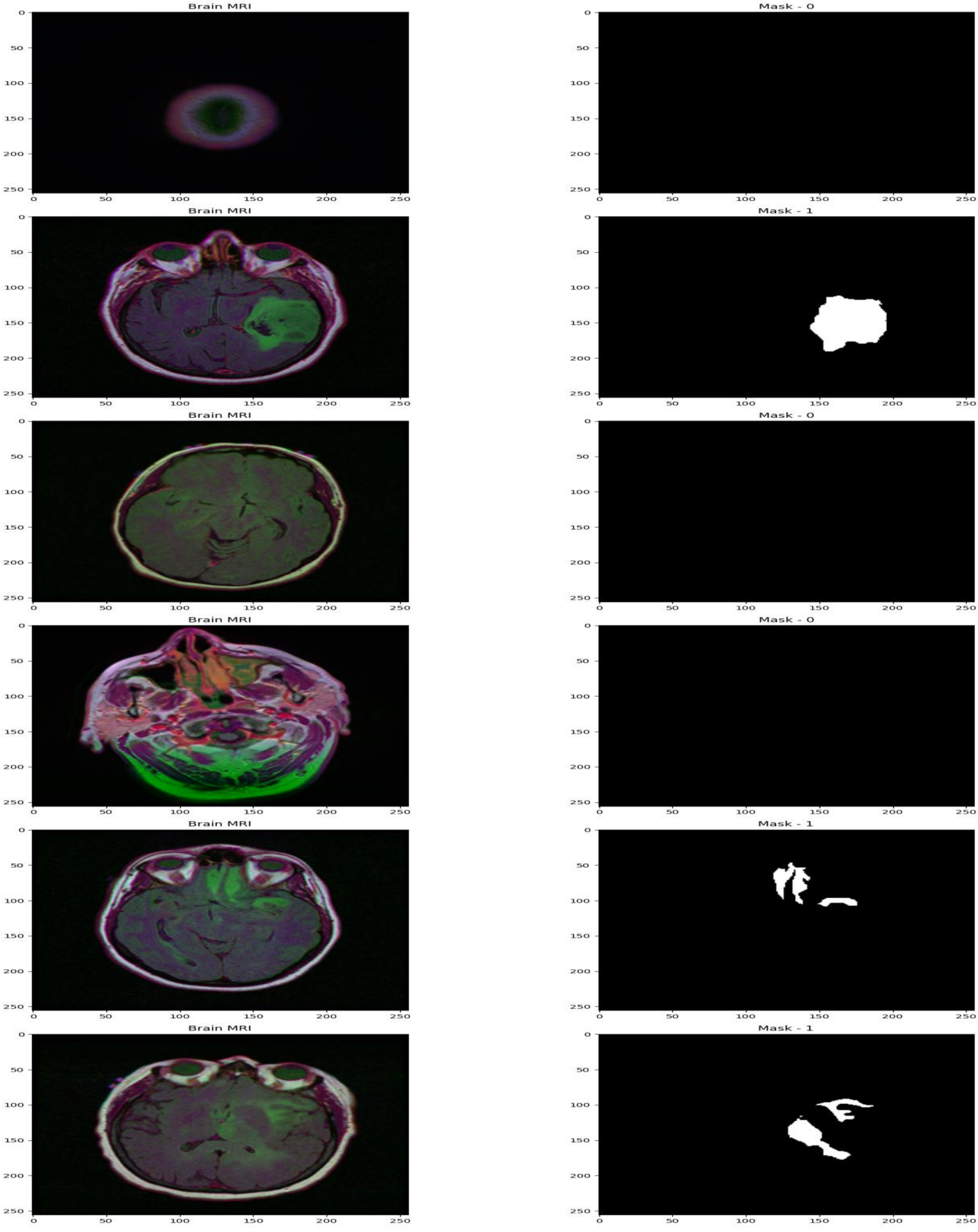


Figure 2. The separation of FLAIR image and Mask image.

3.4. Feature extraction: -

Feature extraction of the image refers to the process that transforms the raw input data into numerical features processed with the preserving the original dataset. Feature extraction [1] decreases the number of resources needed to correctly describe a large set of data. The raw data is converted into the numerical features because the model can't train on the raw data it yields poor results, since the raw data contains lot of information redundancy in it. The Convolutional Layers in the model is responsible for the feature extraction from the image. Convolution layer contains the filters which is also known as the kernels which detect the features like edges, colours throughout the image. If the image is passed to the convolution layer the size of the image will be decreased as well as it will bring the all information in the field together in the single pixel. The group of convolution layer is formed in Sequential manner so that the features of the image was extracted and converted into the smaller size. There are many convolutional layer for eg. 2D convolution, 3D convolution. The 2D Convolution (conv2D) deals the 2 dimensional image where the filters in the convolution slide all over the image, performing an elementwise mathematical operation. There are many mathematical operations can be performed like multiplication, average, nearest neighbour etc. As a result, after this operation, it will have compiled the results in the single pixel based on the pixel location. Same operation will be repeated all over the image, so as the 2D image or 2D matrix of certain features will be transformed into the different 2D images or 2D matrix of features.

3.5. Data visualization: -

Visualization is the process by which one can get the deeper understanding of the distribution of the dataset, training of the model, models prediction etc. Graphical representation or pictorial representation of the structure of the dataset will make the complex problems easier and it will help to make decision making process. Visualization the training of the model will help in the optimization phase. It will be easier to determine which part of the model is valuable and essential. So that optimizer can be designed with the valuable parts and helps to omit some layers so that the models performance was not compromised. Data Visualization will also help to understand the model performance over the training by visualizing it in the graph. It will show the accuracy and loss over every iteration. It will help to take the decision whether the epochs can be adjusted accordingly. Data visualization will be more helpful for the visualization of the predication made by the model. The model will segment the tumour cells after many epochs of training. It will have visualized alongside the actual mask of the image, it will show the difference in the predicted images and true value.

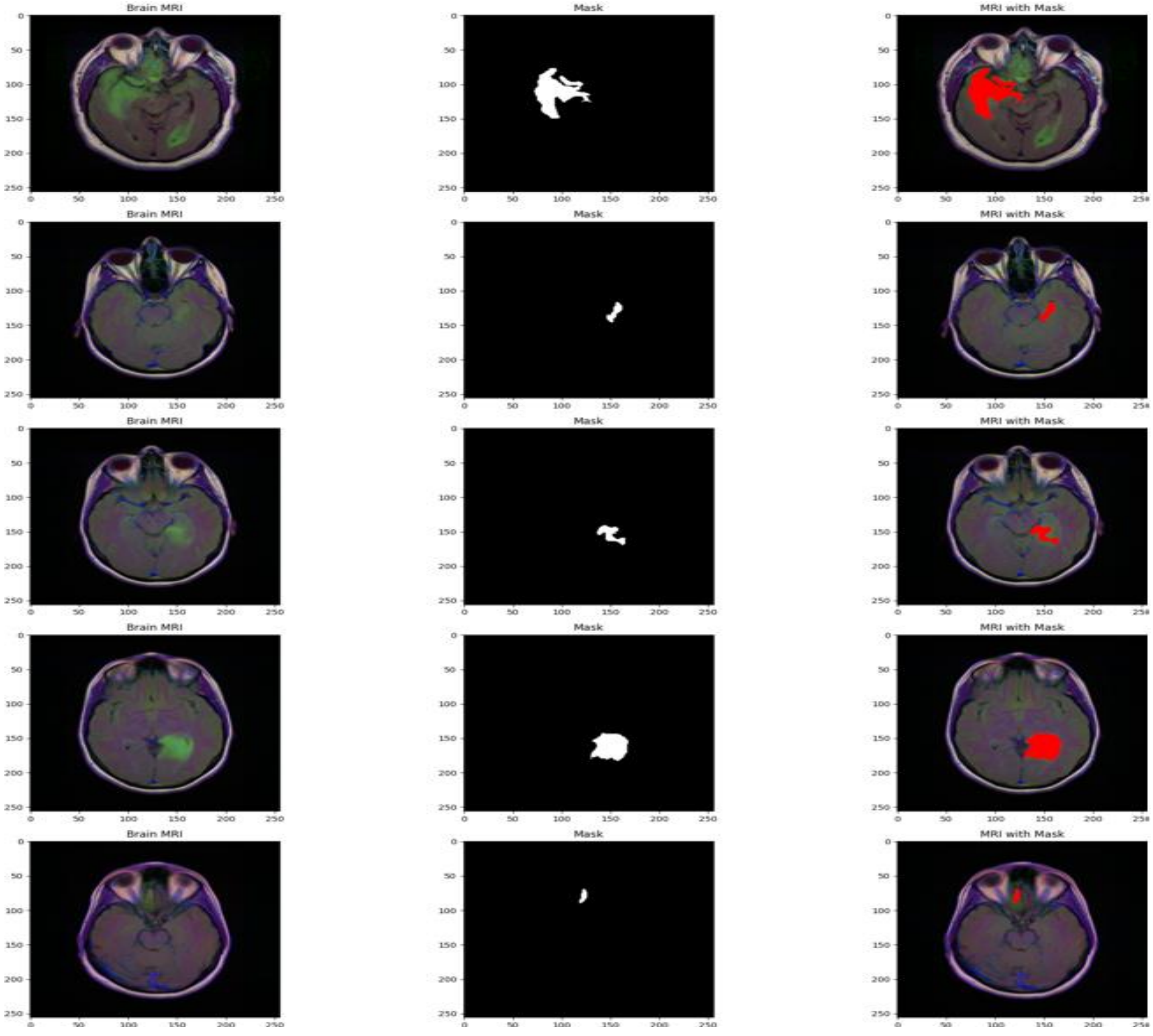


Figure 3: (a) Fluid Attenuated Inversion Recovery (FLAIR) of brain, (b) Mask image of the actual tumour, and (c) the brain's FLAIR image is added with the Mask value.

3.6. Data Augmentation: -

Data Augmentation is the technique to increase the training set by creating the modified copy of the dataset with the help of the already available data. This process will help to increase the size of the dataset and help to balance the variance or diversity in the data. The new dataset was created by transforming the existing data in different forms without losing the existing features. It includes making the small changes to the dataset to create new dataset. The small changes of Geometric transformation include the transform, cropping, rotate, shear, slanting, flipping etc. The Colour Space Transformation is the augmentation method which will change the RGB channels, intensity of the colors and make minor changes in the dataset. Random Erasing in the Augmentation technique will change the intensity or erase the image.

3.7.1. Convolutional Neural Network (CNN)

In Deep learning technique, Convolutional Neural network (CNN) plays an important role. In this project, CNN is used to extract the features from the images and it will classify, compress accordingly. The convolutional neural network is combination of multiple layers of convolutional layers, pooling layers, flattening layer, Dense layer or Fully connected layers.

Convolutional layer in the CNN is important part as it is responsible for the extraction of the main features from the image of multiple layer in sequential manner. A general term for convolution is the grouping in mathematics. The main work of the convolution is to combine two matrices with the mathematical equations to create a third matrices. The Convolutional layer hover over the entire images by moving the Strides and Filters. Filter is responsible for the extraction of the features from the images. The filter will hover over the image by the Stride value. Stride will say that how much pixel does the filter need to be moved further. For example, if the image of **4X4** and the filter size of **2X2** and stride of 1 is applied, then the filter starts from the 1^{st} pixel and it moves to the 2^{nd} pixel and follows till the last pixel of the image. In some cases, applying of filters and strides may results in losing some features in the images. Means, the information in the corners of the image is not extracted. To overcome this problem, the Padding comes into the model. It will add the frame of 0's to the image. So that every nook and corner information of the image is extracted. Every convolutional layer is connected with the activation function. Activation function in added in the convolutional layer is to ensure the non-linearity in the model and to handle and learn more complex relationship in the data. There are many types of Convolutional layer according to the dimension of the data such as conv1D, conv2D, conv3D. In the proposed model of Convolutional neural network there are there are three convolutional layers (conv2D) added in the sequential manner. First convolutional layer with depth of 128. The input is passed in this layer of convolutional layer with the size of **(256,256,3)** which is 3 channel image. The extracted features are passed on to the following sequence of convolutional layer with the depth of 64 and 32. Rectified Linear unit (ReLU) activation function is added in all layers of the convolutions.

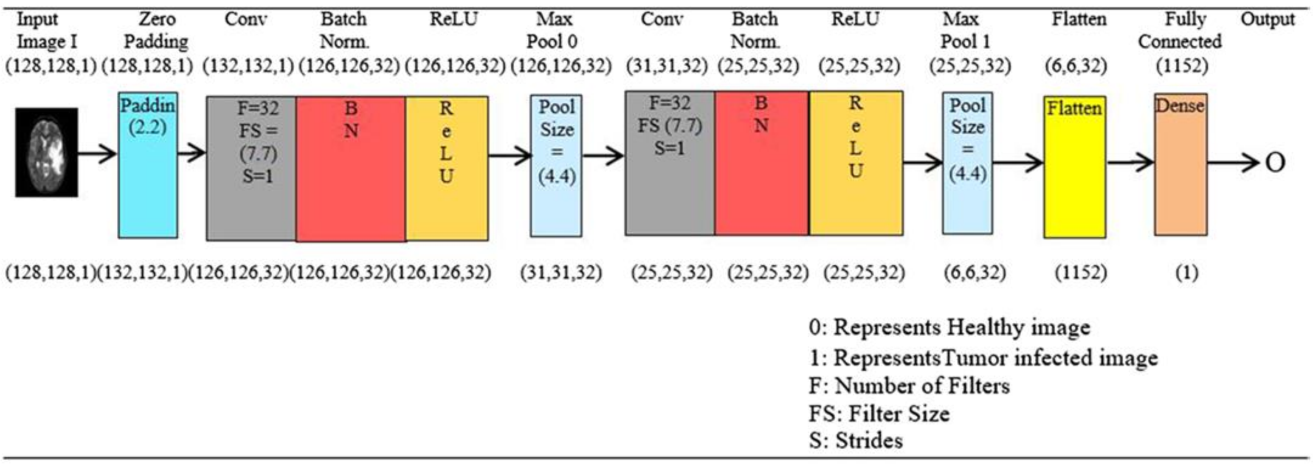


Figure 4: Workflow of CNN. **Ramtekkar et al. (2023)** [1]

Pooling layers in the CNN is responsible for filtering out and reduce the dimension of the extracted features with some mathematical operations from the image. The Pooling layer is responsible for reducing the dimension of the input feature maps. Pooling in general it means that the act of combining two or more things. As the name says it will do some mathematical operations in the input like finding maximum, minimum, average, neighbour and reduce the complexity of the image to the model following. Pooling layers will down samples the volume of neural network by reducing the small translations in the features. There are two types of pooling layers widely used is MaxPooling and AveragePooling. Max pooling is the taking the maximum valued pixel in a particular patch of the feature maps. Average pooling is taking the average among the pixel in a particular patch of the feature map. In the proposed model, Maxpooling2D is used with the size of **2X2**.

Flatten layer act as the bridge between the convolution layer and the dense layer. The responsibility of the flatten layer is to change the shape of the feature maps. The output of the convolutional layer is in the form of multidimensional and the dense layer will accept the one dimensional input. So it makes the multidimensional layered features to one-dimension feature map.

In the architecture of the Convolutional Neural Network (CNN), the last layer is Dense layer or Feed Forward Network. Dense layer is used to classify the images based on the information extracted in the previous layers in the architecture. Dense layer is the combination of multiple neuron connected each other and getting input from the previous layer and perform mathematical operations on the given input before passing it to the next layer or the output layer. Every neuron in the dense layer is added with the activation function to ensure the non-linearity in the data. Rectified Linear unit (RELU) activation is used in the proposed system and the Soft-max activation function is used in the output dense layer because for the binary classification.

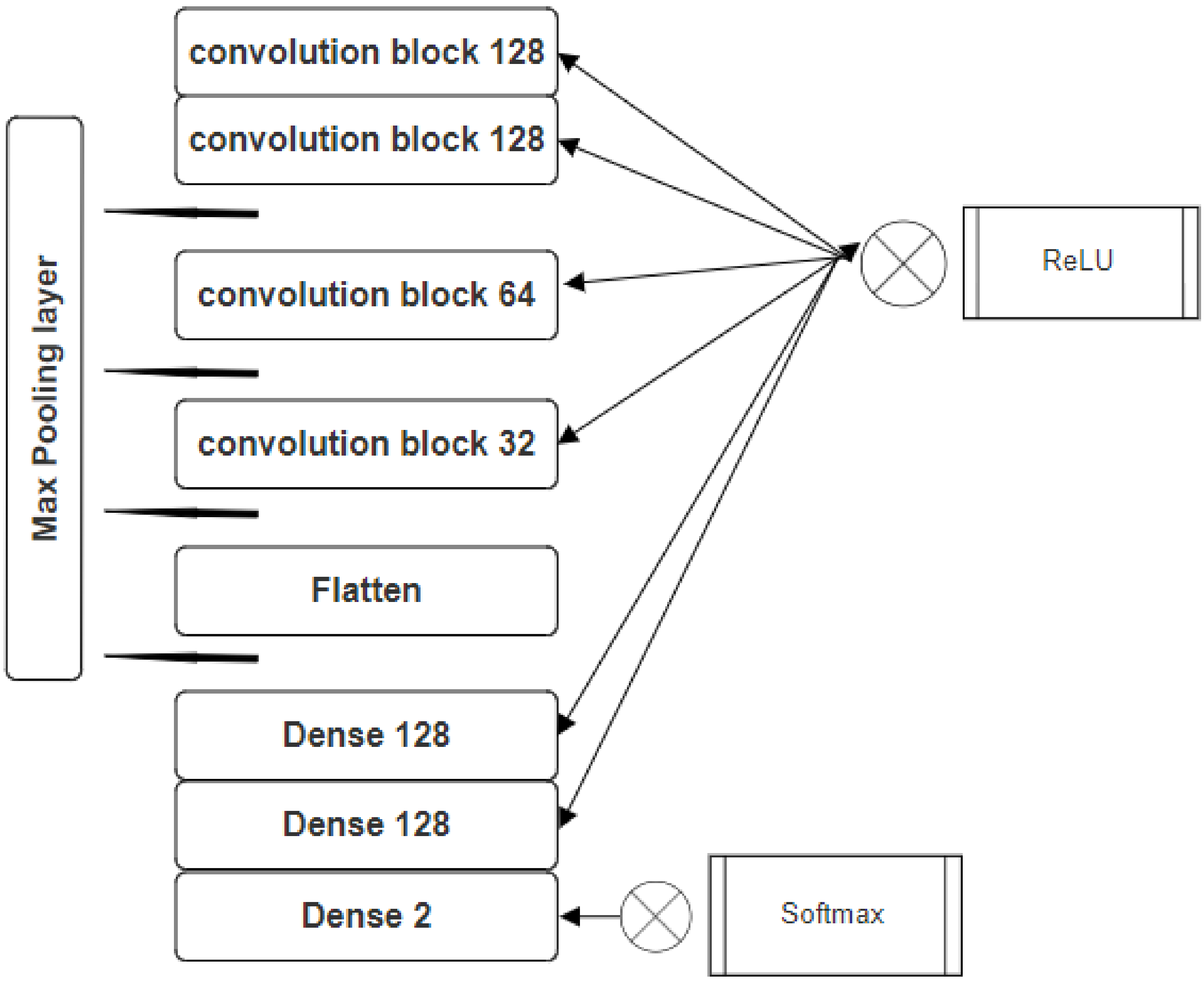


Figure 5. The above flowchart explains the architecture of CNN model.

3.7.2. ResNet50

ResNet50 - residual network and it is a specific form of Convolutional Neural Network (CNN). ResNet50 is a combination of 50 layers [3] of Convolutional Neural Network which is connected in a Sequence of residual block. There are 48 blocks of convolutional layers, one block of Max Pooling layer and one block of Average Pooling layer. These types of residual network are basically a form of artificial neural network (ANN) that forms networks by stacking residual blocks network.

ResNet50 is a powerful image classification model which will be used to train on large number of datasets and to achieve the state-of-art results. ResNet50 allows the models network to train and learn the residual functions that map the input to the desired output. The ResNet50 model will suffer with the problem Vanishing Gradient because the residual network models follow the novel way of building the convolutional neural network (CNN) using the concept called “Shortcut connections”. ResNet50 model is built on the basis of four steps such as the convolutional layers, the identity block, the convolutional block and the fully connected block.

The convolutional layer in the ResNet50 is also responsible for the feature extraction is several layers. The convolutional layers in the ResNet50 consists of several convolutional layers which follows the batch normalization and ReLU activation function. Each layer of the convolutional layer has the ReLU activation function to ensure the non-linearity in the dataset. The Convolutional layers follows the batch normalization, batch normalization can help the model by reduce the sensitivity of the model to the initial weights and converting those weights to easier way to train the model. Usual batch normalization will work like normalizing the output of the previous activation layer by subtracting the mean of the batch and dividing it with the standard deviation of the batch. And then followed up by the max pooling layers, and it is responsible for to reduce the spatial dimension of the feature maps with the assurance of having the most important features of the images.

The identity block and the convolution block in the residual network is responsible is to ensure the dimension of the data that are given to the model. Identity block is the block which is used when there is no difference between the input and the output dimensions of the features. Means there is no problem when the input activation dataset is passed to the model. The dimension of the both input and output activation's is same. The Convolution block is used when there is difference in dimensions between the input activation and the output activation. When there is a difference in the input and the target the model won't train correctly, to overcome this problem convolutional block will alter the dimension of the dataset. These two block is responsible for transferring the datasets to the different blocks.

The fully connected block in the ResNet50 model will perform final classification. The fully connected layer is comprising of the multiple neurons which is well connected with the input layer or the previous layer and the output layers. Based on the extracted features the fully connected block will identify the relation and classify the features.

The proposed ResNet50 model architecture takes the input of **(256,256,3)** dimension dataset. The architecture includes many conv_blocks which holds the convolutional layer, batch normalization, Activation. Those blocks were connected in sequence and each block can get the input and at last all were summed and those were classified by the fully connected layer. The fully connected layer in this architecture are in shape of 3 layers, in which 2 layers where in the shape of **(None, 256)** and the last layer is **(None, 2)** with Soft-max activation function. The total parameters that are passed to this architecture is **25751426**. Out of these the trainable parameters are **25698306** and the testing parameters are **53120**.

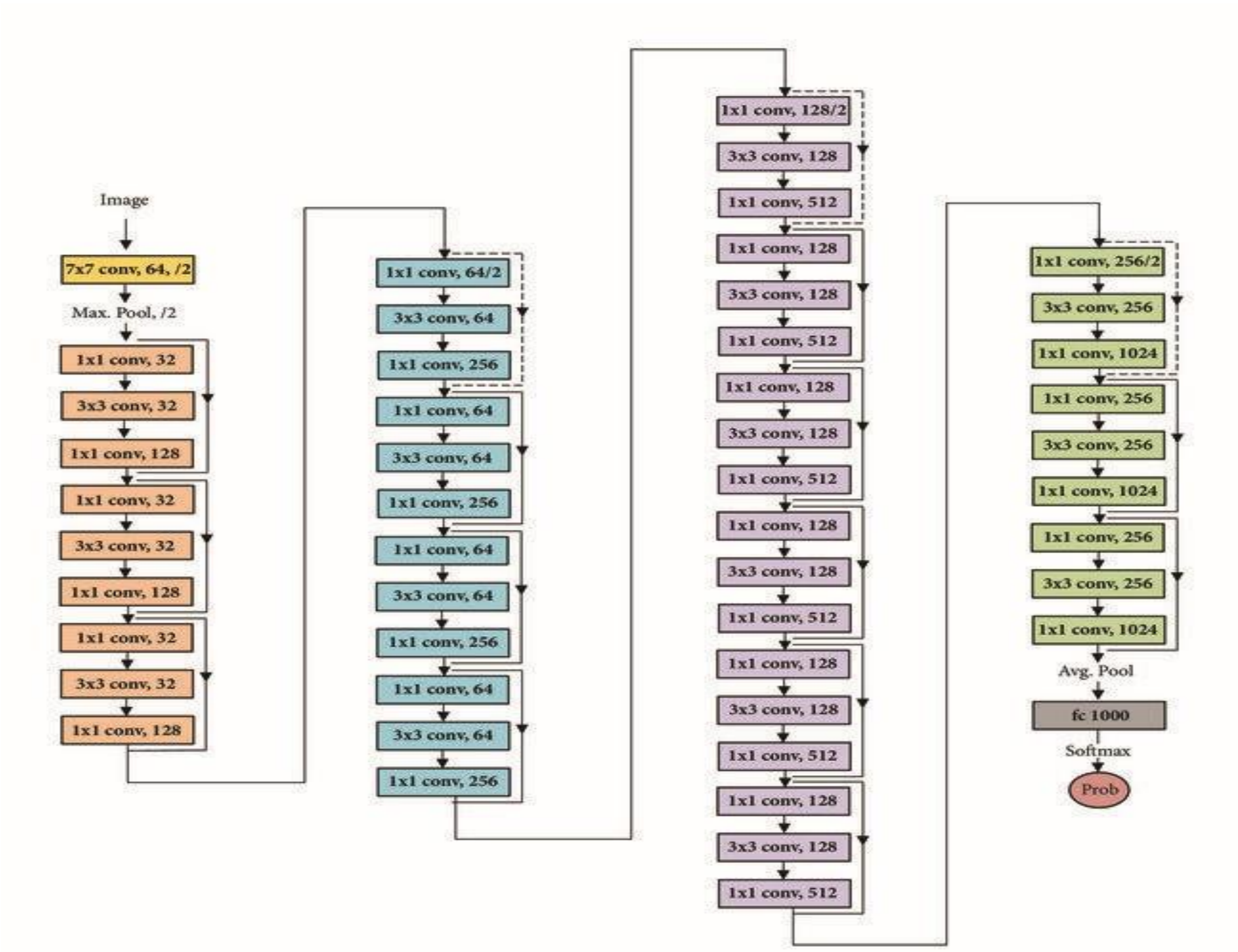


Figure 6. The above diagram explains the block diagram of ResNet50 model. It is cited from "https://www.researchgate.net/publication/326198791_Explicit_Content_Detection_System_An_Approach_towards_a_Safe_and_Ethical_Environment".

3.7.3. Inception V3

The Google developed a model called inception V3 which is the convolutional neural network which is the successor of the GoogleNet. The GoogleNet architecture includes 22 learning layers and nine conceptualization modules [3]. Inception V3 is a part of the Inception project. The main motive of this model Inception V3 is to increase the efficiency and extract the features in the efficient way. With these model the image classification tasks in the Deep Neural Network has been made easier as well very effective.

Inception V3 model is running by using the inception blocks back of it. Inception block is a block which mainly aims to approximate an optimal local sparse structure in a convolutional neural network (CNN). That means inception block consists of multiple sizes of the filter, and it uses different types of filters, it will allow the user to give full access to use multiple filter size instead of restricting the user to stay in single sized filter, in a single block in image, and then it was added together then passed on to the following layer.

Inception V3 architecture are comprised of several layer of convolutional layer which was connected parallel with different filter size, which will help the model to extract the features at different sizes. Inception V3 architecture comprises of usual convolutional layers, pooling layers and fully connected layers. And one more was added in the architecture which is auxiliary classifiers.

In convolutional layers the model uses the factorization technique to reduce the computation cost of the model. It will specifically replace the larger convolutions with a series of smaller convolutions. It will reduce the parameters count and computational complexity of the model. After the convolution layer the model have pooling layers in it. It uses any pooling layers such as max pooling layer and average pooling layer. These pooling layer will down sample the input feature maps. This pooling layers will reduce the spatial dimension of the image with the relevant information.

Auxiliary classifiers are the important layer in the inception model. The Inception V3 model incorporates the auxiliary classifiers in between during the training. The main job of the Auxiliary classifiers is to predict the class labels and it will mitigate the vanishing gradient descent during training by providing some signals to the following layers. The model uses the batch normalization technique for the normalization of the dataset features. Regularization techniques is used to prevent the overfitting and improve generalization performance.

Inception V2 is a Convolutional neural network that is of 48 layers in deep. The proposed model receives the input feature maps in size of **(256,256,3)** dimension dataset. The model consists of convolution layer, activation layer and batch normalization layer connected in parallel for many layers. The total parameters that are passed to the architecture are **21802784**. Out of these the trainable parameters are **21768352** and the testing parameters are **34432**.

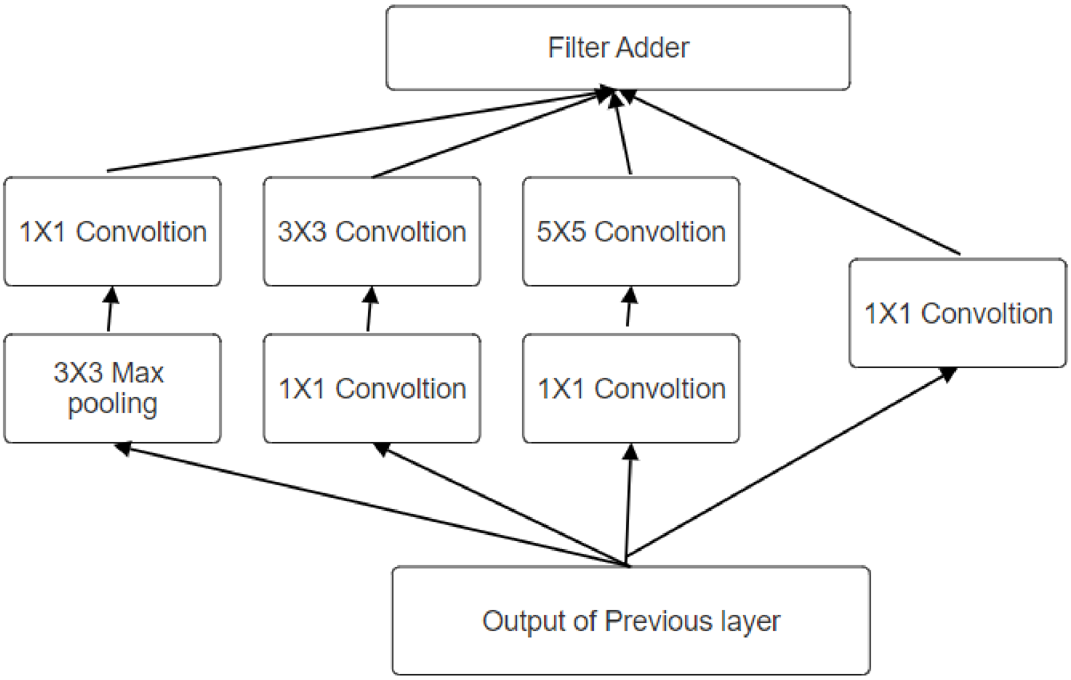


Figure 7. Architecture of Inception block.

3.7.4. Xception

The Xception is a short form of “Extreme Inception” which creates a revolution in the deep learning and computer vision architecture. Xception creates a revolution by extremely reducing the computational cost, input parameters, and it will reduce the time. Xception has a higher accuracy rate over other pre-trained models because xception is lesss susceptible to overfitting.

Regular models use the convolutional layers to extract the features. If the dataset has high spatial channels, then the computational cost, training time of the model will be increased proportionally. As this model will try to extract the information in all three channels parallel, the parameters count is increasing drastically. To overcome this problem Xception model is introduced. This Xception model will be trained with the depthwise convolution layers. This Depthwise convolution layers will work like if there is high spatial channel dataset is passed, then instead of doing all at once it will separately work on every channel and concatenate the channels with **1X1** convolution.

For example, let’s take if the input of the first convolution is 12 and the output mapped to that convolution is 24, the **3X3** channels were passed. The computational parameters of the normal convolutional layer are calculated as **12X24X3X3** which has 2592 parameters. If the same features and dimensions were passed on to the depth wise seperable convolution layer the computational parameters were calculated by **(12X3X3 + 12X24X1X1)** which has 396 parameters.

Depthwise convolution layers is form of convolutional layer which performs separate convolutions on each channel separately of the input features. In detail, that the model will apply the kernel to each channel of the image separately which results in the output channels with the same number of input channels.

The architecture of the Xception model contains Entry flow, Middle flow and Exit flow. The entry flow of the Xception block is of input layer, initial convolutional block, depthwise seperable convolutions (which has depthwise convolutional, pointwise convolutional), residual connection. Each block has the ReLU activation function and pooling layers to ensure the non-linearity and reducing spatial dimensions.

Then the above features are passed to the middle flow which is also of the separable convolutional layer. Rectified linear unit is connected in each layers. This layers were repeated to certain times were it was predefined. Then it was passed to the output flow or the Exit flow. Here the features are passed through ReLU activation, Separable convolutional layer, and at last global average pooling is applied before passing to the fully connected layer. Then the images were classified with the Soft-max classification.

Xception model is a convolutional neural network is 71 layers deep. The proposed xception model receives the input feature maps in (256,256,3) dimension dataset. The model consists of convolution layer, activation layer and batch normalization layer connected in parallel for many layers. The total parameters that are passed to the architecture are **20861480**. Out of these the trainable parameters are **20806952** and the testing parameters are **54528**.

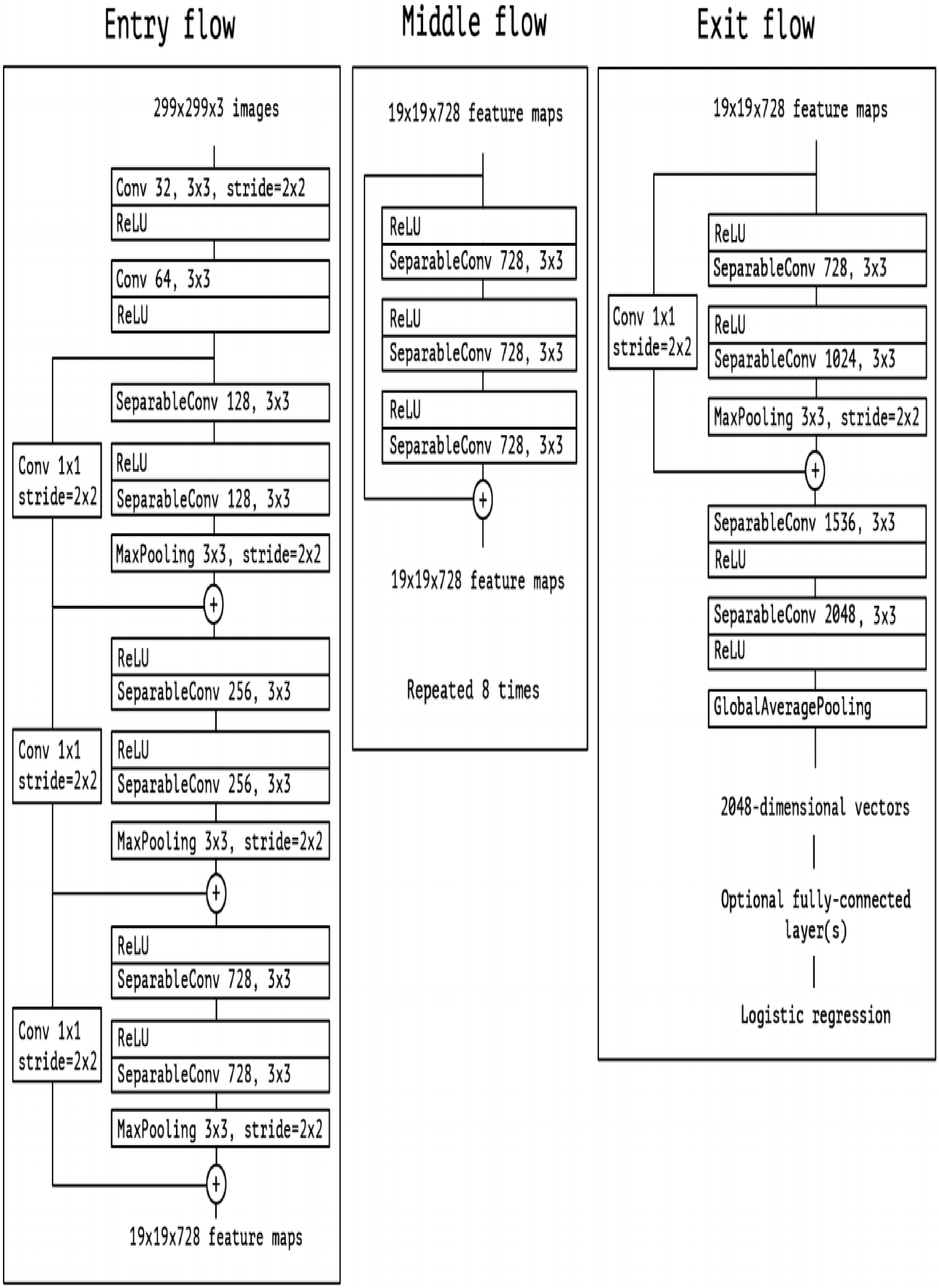


Figure 8. Architecture of Xception model

3.7.5.VGG19

The term VGG in VGG19 stands for Visual Geometry Group, which is a group of convolutional neural network (CNN) architecture. VGG is actually created to increase the depth and the efficiency of the convolutional neural network model. The term 19 in the VGG19 is denoting the depth of the deep - CNN.

VGG19 is a 19 layers of CNN architecture, out of this 16 convolutional layers with 5 layers of max pooling layer and 3 layers of fully connected layer. Convolutional Layer is for the extraction of features from the given data's. Max pooling layer is responsible for the reduction of the dimensional of the features. Fully connected layer is responsible for linear classifier of the image. The model gave the highest accuracy over other models and has the highest accuracy in different models.

Convolutional layer of the model is connected is sequential manner. The depth of the convolutional layer is increasing over the increase in the layers. The layers in the convolutional layer are in the numbers like 64, 128, 256 and 512. The convolutional layer firstly takes the input and it process it with the 2 layers of convolutional layer with 64 nodes. After the extraction of the features in those layer the features are passed to the max pooling layer, here the dimension of the feature cut down for the simplicity to the model. Then it was passed to 2 layers of convolutional layer with 128 layers and then Max pooling layer. And then followed by 4 layers of 256 convolutional layer, max pooling layer, 8 layers of convolutional layer with 512 nodes. Every convolutional layer nodes are connected with the ReLU activation function for the non-linearity in the data.

The fully connected layer in the block is responsible for the classification of the image with the extracted image features in the consecutive layers. There are 3 layers of fully connected layers are connected in the sequence to train on the features. First two layer of the dense layer contains the 4098 neurons, and lastly with the 1000 neurons to train on the features. Soft-max activation function connected to the last layer to classification.

VGG19 model is a convolutional neural network is 16 layers deep. The proposed xception model receives the input feature maps in (256,256,3) dimension dataset. The model consists of convolution layer, activation layer and batch normalization layer connected in parallel for many layers. The total parameters that are passed to the architecture are **20024384.** Out of these the trainable parameters are **20024384** and the testing parameters is **0**.

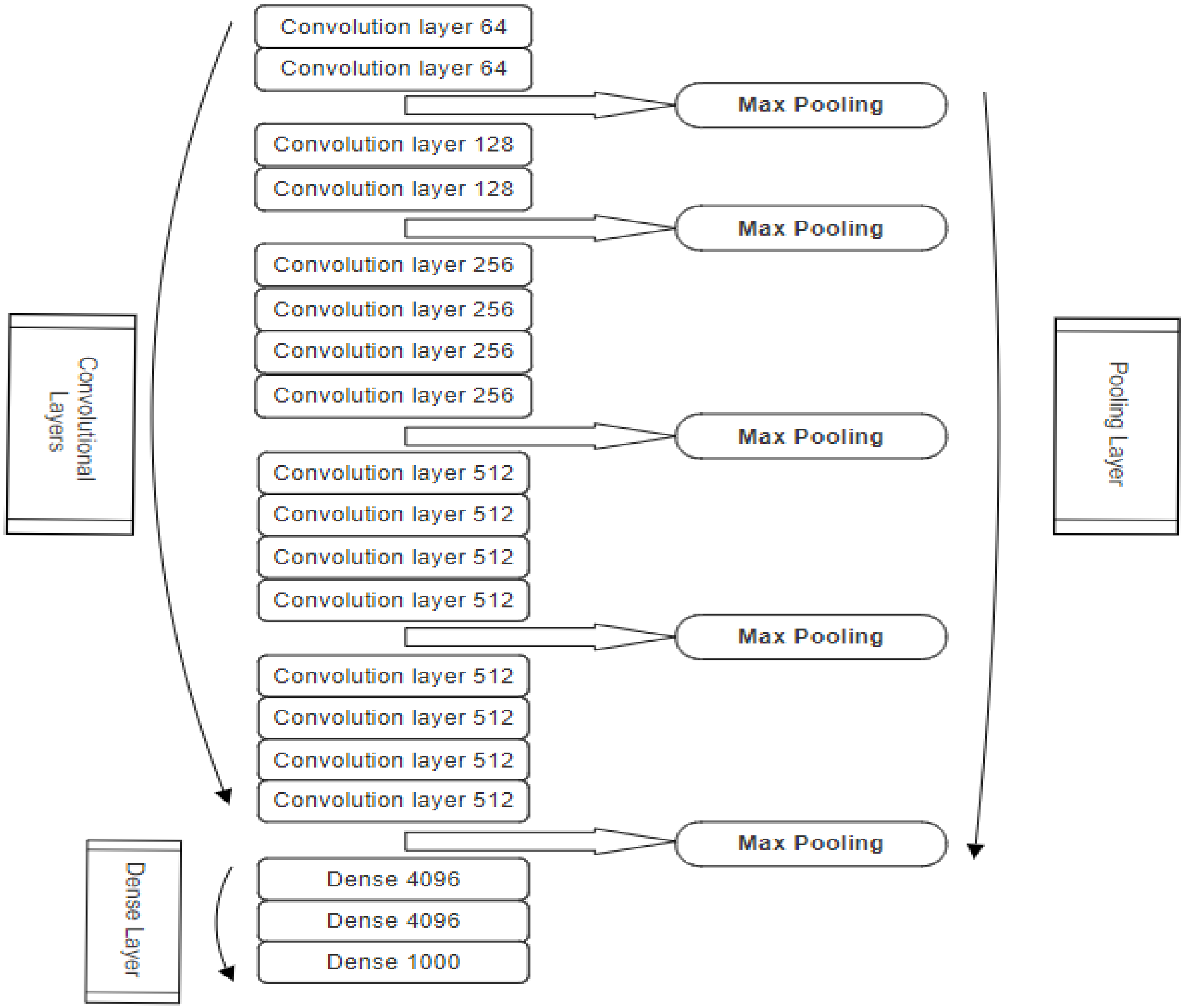


Figure 9. The above diagram represents the architecture of VGG19 model

3.7.6. MobileNetV2

MobileNetV2 is an architecture of convolutional neural network which works well on the mobile devices. The architecture is quite opposite to the residual architecture. The residual blocks is bottleneck shape with the input and output layers were thick and the middle layers are thin. In MobileNetV2 the architecture is inverted residual with the input and the output of the blocks are thin layer whereas the middle layer is thicker. The middle layer expansion uses the lightweight depth-wise convolutions to filter feature source of non-linearity.

The MobileNetV2 model introduces two hyper-parameters, one is Width multiplier and the other is Resolution multiplier. The reason to introduce this hyper-parameter is to achieve the balance between the complexity and accuracy. The width multiplier's work is to scale and adjust the number of channels in each layer, which is always a trade-off between the model size and the performance. Whereas the work of the resolution multiplier is to reduce the input resolution of the network, and even reduces the computational requirements while maintain the performance to some extent.

The MobileNetV2 is a 53 layer of deep lightweight CNN model. The model takes the input image of shape of (256,256,3). The architecture is formed with the inverted bottleneck blocks which contains depth-wise convolutions. These inverted blocks were connected in sequential form and in every layer the feature maps were encoded to low dimensional subspaces. The depth-wise separable convolution is used to trim-down the redundant information. The model consists of convolution layer, activation layer and batch normalization layer connected in parallel for many layers. The total parameters that are passed to the architecture are **20024384.** Out of these the trainable parameters are **20024384** and the testing parameters is **0**.

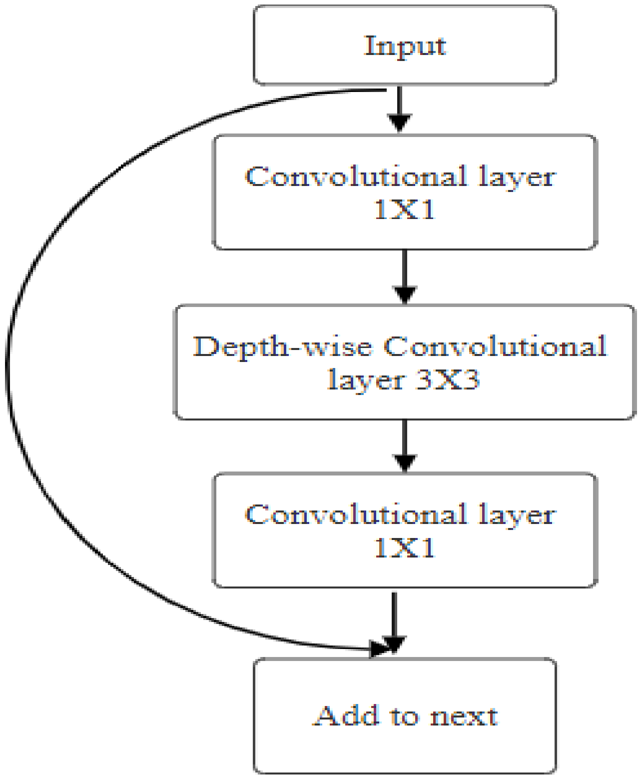
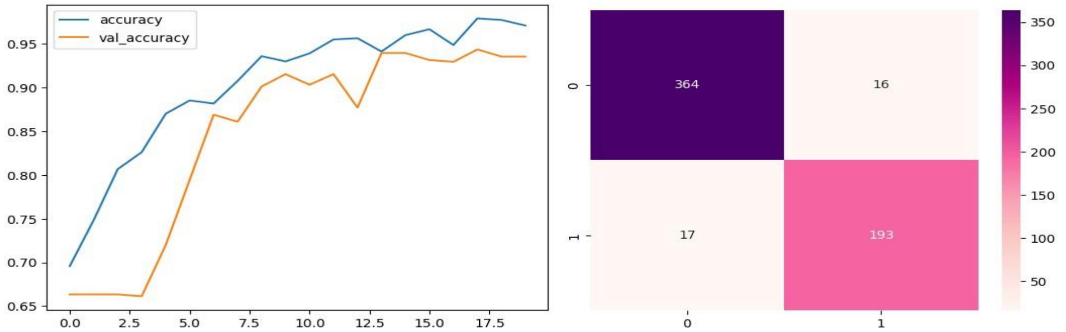


Figure 10. The above architecture represents the architecture of MobileNetV2.

3.8. Results

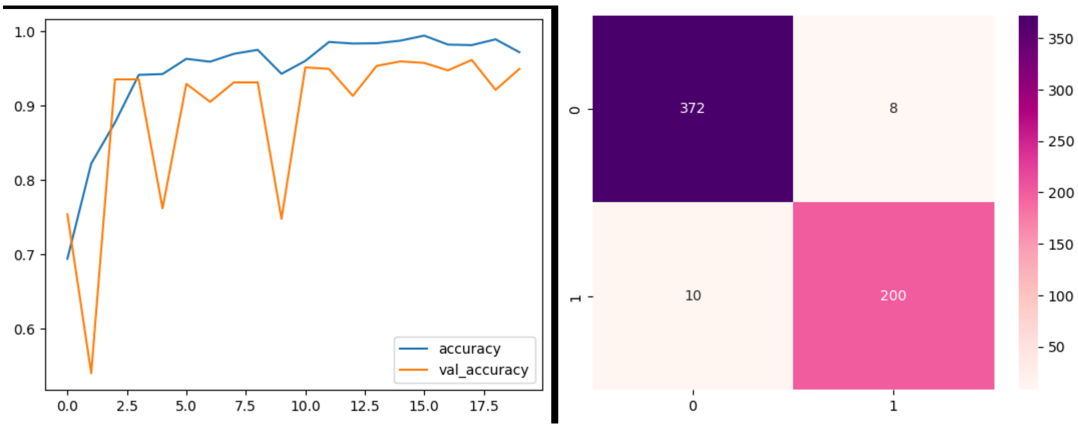
The dataset for the brain tumour is pre-processed, segmented, extracted, augmented and the deep learning models were introduced for the dataset. The deep learning models were included into this architecture and run successfully. The pre-trained models like ResNet50, Xception, Inceptionv3, VGG19, MobileNetV2 were produced with the good accuracy of 94.4%, 96.9%, 94.23%, 65%, 64%. The validation accuracy and the accuracy is analysed and plotted in a graph. The confusion matrix is calculated with the precision, recall, F1 score. From the below shown graph and in the scatter matrix it is quite evident that the model performs and trained well in the model.

Top 3 pre-trained model results and performance.



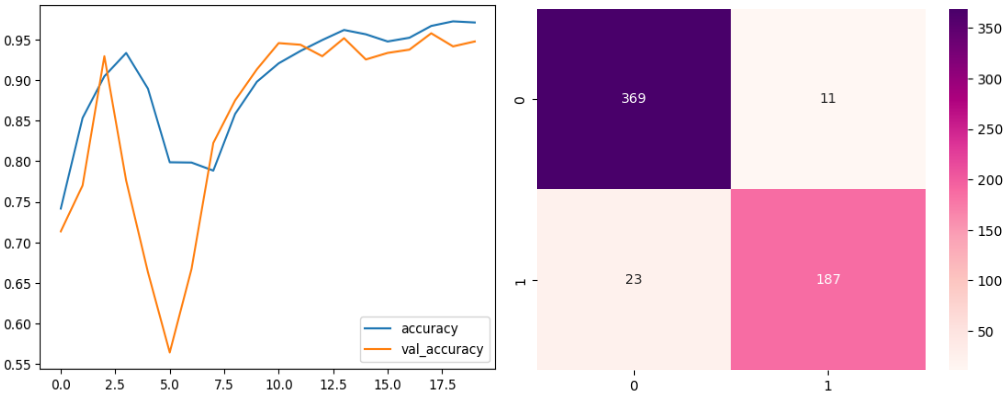
	precision	recall	f1-score	support
0	0.96	0.96	0.96	380
1	0.92	0.92	0.92	210
accuracy			0.94	590
macro avg	0.94	0.94	0.94	590
weighted avg	0.94	0.94	0.94	590

Figure 11.1. ResNet50 model accuracy chart, confusion matrix and classification report.



	precision	recall	f1-score	support
0	0.97	0.98	0.98	380
1	0.96	0.95	0.96	210
accuracy			0.97	590
macro avg	0.97	0.97	0.97	590
weighted avg	0.97	0.97	0.97	590

Figure 11.2. Xception model accuracy chart, confusion matrix and classification report.




```
              precision    recall  f1-score   support

           0       0.94      0.97      0.96       380
           1       0.94      0.89      0.92       210

    accuracy                           0.94       590
   macro avg       0.94      0.93      0.94       590
weighted avg       0.94      0.94      0.94       590
```

Figure 11.3. InceptionV3 model accuracy chart, confusion matrix and classification report.

S.no	Models	Accuracy
1	ResNet50	94.4067
2	Xception	96.9491
3	InceptionV3	94.2372
4	VGG19	64.4067
5	MobileNetV2	62.0338

Table 1. Result of the pre-trained model.

4. Conclusion

The design and development of the DL model for the dataset's classification of brain tumours is the study's primary goal. The task of classifying brains into benign and malignant classes is difficult because of the substantial amount of complexity in the dataset. The suggested models have been designed to function with a comparatively high amount of picture data. The work provides demonstrations for segmentation, feature extraction, and pre-processing of data. The visualization and augmentation of the dataset are demonstrated. The dataset was explored using the pre-trained models, and the model accurately predicted the location of the tumour. For the purpose of to detect brain tumours, five robust deep learning pre-trained models were used. Good accuracy was obtained from the pre-trained models ResNet50, Xception, Inceptionv3, VGG19, and MobileNetV2. With a 96.94% accuracy rate, the InceptionV3 model yielded excellent outcomes for the TCGA dataset. The study indicated the model with the best accuracy provider among the ones listed above. Compared to other models, this one is better and has a higher classification accuracy. The plan is to build on this work in the future to further improve the models' accuracy in performance.

Reference: -

1. Praveen Kumar Ramtekkar, Anjana Pandey, Mahesh Kumar Pawar (2023) Accurate detection of brain tumour using optimized feature selection based on deep learning techniques. Multimedia Tools and Applications (2023) 82:44623–44653. https://doi.org/10.1007/s11042-023-15239-7.
2. Sandeep Kumar, Shilpa Choudhary, Arpit Jain, Karan Singh, Ali Ahmadian, Mohd Yazid Bajuri (2023) Brain Tumour Classification Using Deep Neural Network and Transfer Learning. Brain Topography (2023) 36:305–318. https://doi.org/10.1007/s10548-023-00953-0.
3. Shaimaa E. Nassar, Ibrahim Yasser, Hanan M. Amer, Mohamed A. Mohamed (2024) A robust MRI-based brain tumour classification via a hybrid deep learning technique. The Journal of Supercomputing (2024) 80:2403–2427. https://doi.org/10.1007/s11227-023-05549-w.

4. Saqlain Razzaq, Muhammad Adeel Asghar, Abdul Wakeel, Muhammad Bilal (2024) Least complex oLSVN-based computer-aided healthcare system for brain tumour detection using MRI images. Journal of Ambient Intelligence and Humanized Computing (2024) 15:683–695. https://doi.org/10.1007/s12652-023-04725-3.
5. Ejaz Ul Haq, Huang Jianjun, Kang Li, Hafeez Ul Haq, Tijiang Zhang (2023) An MRI-based deep learning approach for efficient classification of brain tumours. Journal of Ambient Intelligence and Humanized Computing (2023) 14:6697–6718. https://doi.org/10.1007/s12652-021-03535-9.
6. P. Ramya, M. S. Thanabal, C. Dharmaraja (2021) Brain tumour segmentation using cluster ensemble and deep super learner for classification of MRI. Journal of Ambient Intelligence and Humanized Computing (2021) 12:9939–9952. https://doi.org/10.1007/s12652-021-03390-8.
7. Maibam Mangalleibi Chanu, Khelchandra Thongam (2021) Computer-aided detection of brain tumour from magnetic resonance images using deep learning network. Journal of Ambient Intelligence and Humanized Computing (2021) 12:6911–6922. https://doi.org/10.1007/s12652-020-02336-w.
8. B. J. D. Kalyani, K. Meena, E.Murali, L. Jayakumar, D. Saravanan (2023) Analysis of MRI brain tumour images using deep learning techniques. Soft Computing (2023) 27:7535–7542. https://doi.org/10.1007/s00500-023-07921-7.
9. Rammurthy D, Mahesh PK (2020) Whale Harris Hawks optimization based deep learning classifier for brain tumour detection using MRI images. J King Saud Univ Comput Inf Sci 1–14.
10. Sohaib Asif, Ming Zhao, Fengxiao Tang, Yusen zhu (2023) An enhanced deep learning method for multi-class brain tumour classification using deep transfer learning. Multimedia Tools and Applications (2023) 82:31709-31736.\
11. Baiju Babu Vimala, Saravanan Srinivasan, Sandeep Kumar Mathivanan, Mahalakshmi, Prabhu Jayagopal, Gemmachis Teshite Dalu (2023) Detection and classification of brain tumour using hybrid deep learning models. Scientific Reports 13:23029. https://doi.org/10.1038/s41598-023-50505-6.
12. Abhishek Anil, Aditya Raj, H Aravind Sarma, Naveen Chandran R, Deepa P L (2019) Brain Tumour detection from brain MRI using Deep Learning. ISSN(Online) : 2456-8910.
13. Chirodip Lodh Choudhury, Chandrakanta Mahanty, Raghvendra Kumar, Brojo Kishore Mishra (2020) Brain Tumour Detection and Classification Using Convolutional Neural Network and Deep Neural Network. IEEE Xplore.
14. Heba Mohsen, El-Sayed A. El-Dahshan, El-Sayed M. El-Horbaty, Abdel-Badeeh M. Salem (2018) Classification using deep learning neural networks for brain tumours. Future Computing and Informatics Journal 3 (2018) 68e71.
15. Javeria Amin, Muhammad Sharif, Mudassar Raza, Mussarat Yasmin (2018) "Detection of Brain Tumour based on Features Fusion and Machine Learning".
16. Sidra Sajid, Saddam Hussain, Amna Sarwar (2019) "Brain TumourDetection and Segmentationin MR Images Using Deep Learning".
17. Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do, Kaori Togashi (2018) "Convolutional neural networks: an overview and application in radiology"
18. B Kokila, M S Devadharshini, A Anitha and S Abisheak Sankar (2021) "Brain Tumour Detection and Classification Using Deep Learning Techniques based on MRI Images"
19. P Gokila Brindha, M Kavinraj, P Manivasakam, P Prasanth (2020) " Brain tumour detection from MRI images using deep learning techniques".
20. G.Hemanth, M.Janardhan, L.Sujihelen (2019) " DESIGN AND IMPLEMENTING BRAIN TUMOUR DETECTION USING MACHINE LEARNING APPROACH ".
21. Siar M, Teshnehlab M (2019) "Brain Tumour Detection Using Deep Neural Network and Machine Learning Algorithm". 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE).
22. Grampurohit S, Shalavadi V, Dhotargavi V. R, Kudari M, Jolad S. (2020) "Brain Tumour Detection Using Deep Learning Models". 2020 IEEE India Council International Subsections Conference (INDISCON).
23. Choudhury C. L, Mahanty C, Kumar R, Mishra B. K (2020) "Brain Tumour Detection and Classification Using Convolutional Neural Network and Deep Neural Network". 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA).
24. Methil A. S (2021) "Brain Tumour Detection using Deep Learning and Image Processing". 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS).
25. Saleh A, Sukaik R, Abu-Naser S. S. (2020) "Brain Tumour Classification Using Deep Learning". 2020 International Conference on Assistive and Rehabilitation Technologies (iCareTech).
26. Nair V, Hinton GE (2010) Rectified linear units improve restricted Boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning.
27. Ramachandran P, Zoph B, Le QV (2017) Searching for activation functions.
28. Glorot X, Bordes A, Bengio Y (2011) Deep sparse rectifier neural networks. In: Proceedings of the 14th International Conference of Artificial Intelligence and Statistics.
29. Mehrotra R, Ansari MA, Agrawal R, Anand RS (2020) "A transfer learning approach for AI-based classification of brain tumours".
30. Nitish Z, Pawar V (2012) "GLCM textural features for brain tumour classification".
31. Ouseph NC, Shruti K (2017) "A reliable method for brain tumour detection using CNN technique".
32. Rani S, Ghai D, Kumar S, Kantipudi MP, Alharbi AH, Ullah MA (2022) "Efficient 3D AlexNet architecture for object recognition using syntactic patterns from medical images".